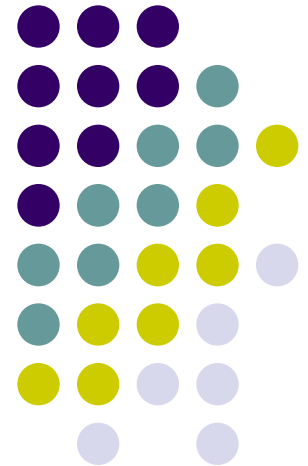
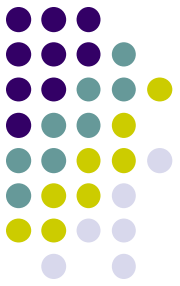


Optimization: Linear programming

Hongxin Zhang

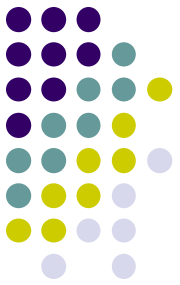
2020-03-31





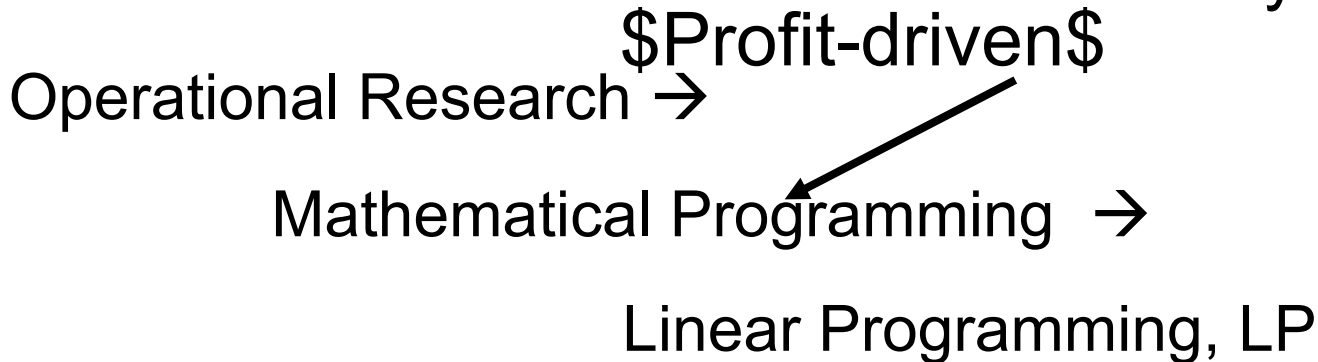
Content

- Linear programming
- Non-linear optimization
- Text books
 - 《线性规划》
 - 张建中，许绍吉，科学出版社
 - 《最优化理论与方法》
 - 袁亚湘，孙文瑜，科学出版社
 - 《数学规划》
 - 黄红选，韩继业，清华大学出版社



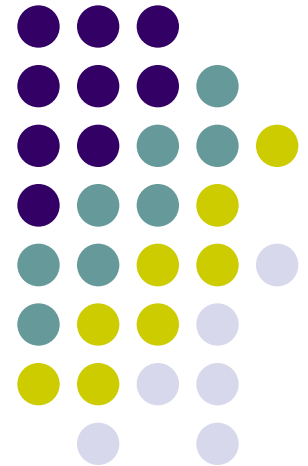
Linear programming

- In the production practice, we often need to consider how to make full use of available resources for maximum economic efficiency.

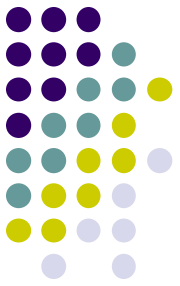


- In 1947, G. B. Dantzig proposed the simplex method to solve linear programming problem.
- Linear programming are more widely used since current computer is able to deal with complex linear programming problem, which has thousands of constraints and decision variables.

Basic Theory

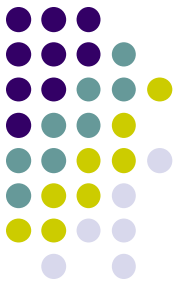


Problem definition (linear programming, LP)



- A mobile game company is going to develop and operate large mobile games. Now they have three technical teams A,B,C.
 - RPG game system
 - Gain 40 million profit per game
 - Need team A cooperate with team B
 - Team A spends 2 month and team B spends 1 month;
 - Social mobile game
 - Gain 30 million profit per game
 - Need cooperation among team A,B, C
 - Every team spends 1 month
 - In 2019~2020, team A can work 10 month, team B 8 month, team C 7 month.
- Question: What is the best strategy to gain maximum profit?

Problem definition (linear programming, LP)



- Suppose that develop x_1 RPG games, x_2 social mobile games
- Objective function

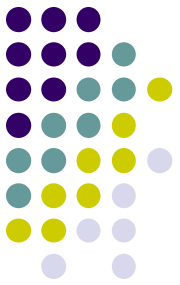
$$\max \quad z = 4x_1 + 3x_2 \quad \left\{ \begin{array}{l} 2x_1 + x_2 \leq 10 \\ x_1 + x_2 \leq 8 \end{array} \right.$$

- Constraint conditions $\left\{ \begin{array}{l} x_2 \leq 7 \\ x_1, x_2 \geq 0 \end{array} \right.$

Additional conditions :
 x_1 and x_2 are integer.

Discussion? ! ?

Problem definition (linear programming, LP)



- Minimize or maximize a linear cost function subject to a set of linear equality and inequality constraints.

- Formal definition:

$$\min c_1x_1 + \dots + c_nx_n$$

s. t.

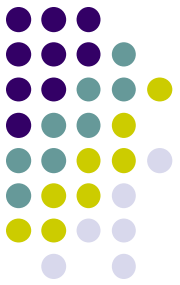
$$a_{11}x_1 + \dots + a_{1n}x_n \leq b_1$$

...

$$a_{m1}x_1 + \dots + a_{mn}x_n \leq b_m$$

$$x_1, \dots, x_n \geq 0$$

Problem definition (linear programming, LP)



- Minimize or maximize a linear cost function subject to a set of linear equality and inequality constraints.

- Formal definition :

$$\min \mathbf{c}^T \mathbf{x}$$

$$\text{s.t. } \mathbf{A}\mathbf{x} \leq \mathbf{b}$$

$$\mathbf{x} \geq \mathbf{0}$$

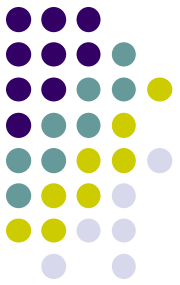
objective function

c: cost vector

A: constraint matrix

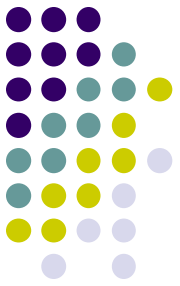
b: right-hand-side vector

Problem definition (linear programming, LP)



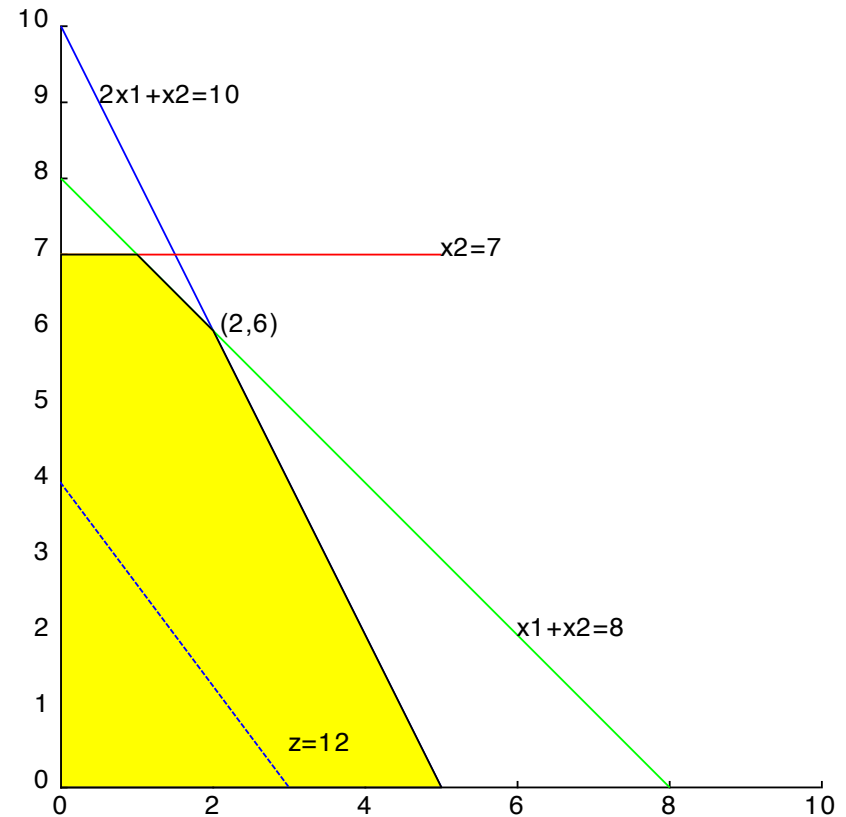
- **X**: A vector satisfies all of the constraints, called a feasible solution or feasible point
- **D**: A set of all feasible points, called feasible region
- **LP problem**:
 - $D = \emptyset$, no solution or infeasible
 - $D \neq \emptyset$, but the objective function is boundless in D : boundless
 - $D \neq \emptyset$, and the objective function has finite optimal solution: has optimal solution

Graphical solution of a linear programming problem

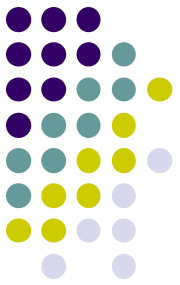


$$\max \quad z = 4x_1 + 3x_2$$

$$\begin{cases} 2x_1 + x_2 \leq 10 \\ x_1 + x_2 \leq 8 \\ x_2 \leq 7 \\ x_1, x_2 \geq 0 \end{cases}$$



Implementation of linear programming in Matlab

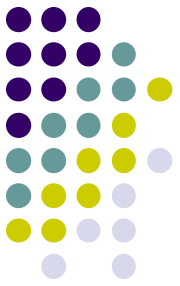


- Matlab specify the standard form of linear programming as:

$$\min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} \quad \text{s. t.} \quad \mathbf{Ax} \leq \mathbf{b}$$

- Including:
 - \mathbf{C} , \mathbf{x} are *n-dimensional* column vectors,
 - \mathbf{B} is a *m-dimensional* column vector,
 - \mathbf{A} is a *m×n* matrix.
- <http://www.mathworks.cn/cn/help/optim/ug/linprog.html>

Implementation of linear programming in matlab



- Basic functional form :

`linprog(c, A, b)`

- Return the value of vector x

- Example:

`[x,fval]=linprog(c, A, b, Aeq, beq, LB, UB, X0, OPTIONS)`

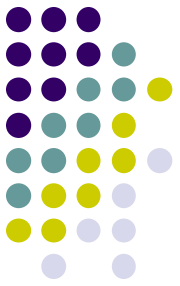
`fval` returns the value of the objective function ,

`Aeq`, `beq` correspond to the equality constraints, `Aeq X = beq`

`LB`, `UB` are lower and upper bounds of the variable x ,
respectively. `X0` is the starting point of X

`OPTIONS` is a structure of control parameters.

Implementation of linear programming in Matlab



- Example:

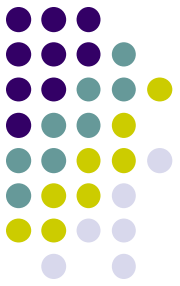
$$\max z = 2x_1 + 3x_2 - 5x_3$$

$$\begin{cases} x_1 + x_2 + x_3 = 7 \\ 2x_1 - 5x_2 + x_3 \geq 10 \\ x_1, x_2, x_3 \geq 0 \end{cases}$$

- Matlab code:

```
c=[2,3,-5];  
a=[-2,5,-1];  
b=-10;  
aeq=[1,1,1];  
beq=7;  
x=linprog(  
    -c,a,b,aeq,beq,  
    zeros(3,1))  
value=c'*x
```

Implementation of linear programming in Python



- CVXOPT :
- <http://abel.ee.ucla.edu/cvxopt/index.html>

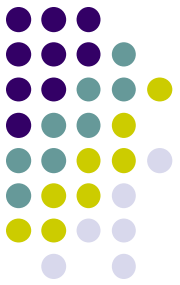


Standard LP problem

- In practice, the LP form we get is quite different:
 - Objective function is a maximization problem or minimization problem
 - Constraints are a set of linear equality and inequality constraints
 - Variables have upper bound or lower bound or no bound

$$\begin{aligned}\min z &= \mathbf{c}^T \mathbf{x} \\ \text{s.t. } A\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0}\end{aligned}$$

Standard form LP problem



$$\begin{aligned} \min z &= \mathbf{c}^T \mathbf{x} \\ \text{s.t. } A\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0} \end{aligned}$$

- Standardization

- Objective function transformation $\max z \rightarrow \min (-z)$
- Constraints transformation (introduce slack variable)

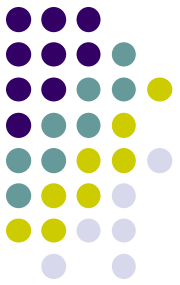
$$\sum_{j=1}^n a_{ij}x_j \leq b_i \Leftrightarrow \begin{cases} \sum_{j=1}^n a_{ij}x_j + x_{n+i} = b_i \\ x_{n+i} \geq 0 \end{cases}$$

- Variables nonnegative constraints

$$x_j \geq l_j \Leftrightarrow y_j \geq 0, y_j = x_j - l_j$$

$$x_j \text{ free variable} \Leftrightarrow u_j \geq 0, v_j \geq 0, x_j = u_j - v_j$$

Standard form LP problem



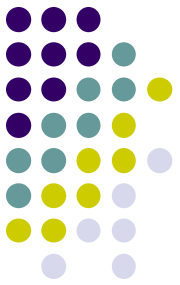
$$\begin{aligned}\min z &= \mathbf{c}^T \mathbf{x} \\ \text{s.t. } A\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0}\end{aligned}$$

- Standardization

- Objective function transformation $\min |x_1| + |x_2|?$
- Variable substitution

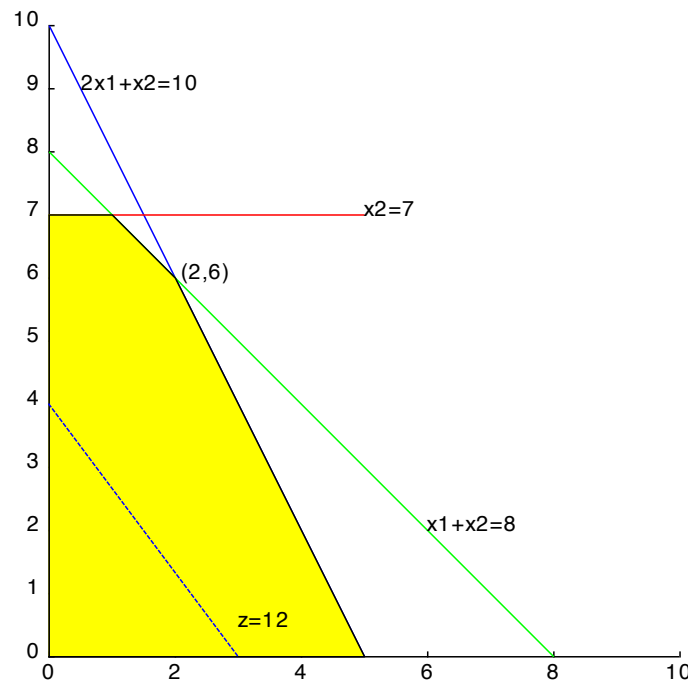
$$x_j \text{ free variable} \Leftrightarrow x_j^- \geq 0, x_j^+ \geq 0$$

$$z = \sum_j |x_j| = \sum_j (x_j^+ - x_j^-)$$

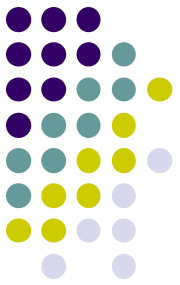


Feasible region D

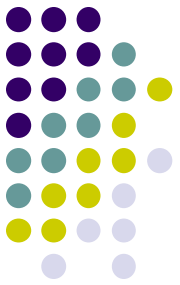
- Discuss the structure of feasible region $D = \{ \mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \}$
 - First, discuss the set $K = \{ \mathbf{x} \mid A\mathbf{x} = \mathbf{b} \}$



Affine set



- Affine set: A set $S \subseteq E^n$, for any $\mathbf{x}, \mathbf{y} \in S$, $\lambda \in E^1$, we have $\lambda \mathbf{x} + (1-\lambda) \mathbf{y} \in S$. Then we call S an affine set.
- Example:
 - Both line and plane of E^3 are affine sets, and empty set \emptyset is also an affine set
 - The sufficient and necessary condition that S is a subset of E^n is that S is an affine set containing the origin
 - Every non-empty affine set S is parallel to a unique subspace L
 - Set shift: $S + \mathbf{p} = \{ \mathbf{x} + \mathbf{p} \mid \mathbf{x} \in S \}$
 - The dimensions of a non-empty affine set S is defined as those of its parallel subspace, $\dim(S)$



Affine set

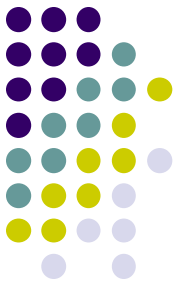
- Affine sets with dimensions 0,1,2 are point, line, plane, respectively.
- An affine set in E^n with $n-1$ dimensions is called hyperplane
- Given $b \in E^1$ and nonzero vector $\mathbf{a} \in E^n$, then the set

$$H = \{ \mathbf{x} \mid \mathbf{a}^T \mathbf{x} = b \}$$

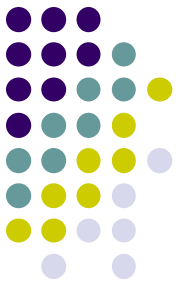
is hyperplane in E^n , and any hyperplane in E^n can be expressed as this form.

- \mathbf{a} is a normal vector of H
- Fix \mathbf{a} , b varies, $\mathbf{a}^T \mathbf{x} = b$ represent a set of hyperplane in E^n

Affine set



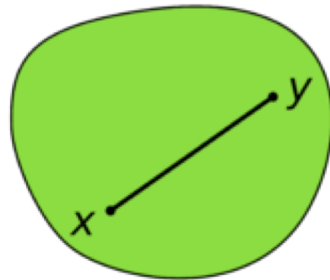
- Properties of general affine sets:
 - Given $m \times n$ real matrix A and $\mathbf{b} \in E^m$, then $K = \{\mathbf{x} \in E^n \mid A\mathbf{x} = \mathbf{b}\}$ is an affine set in E^n , and any affine set in E^n can be expressed as this form.
- The set $K = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}\}$ is an affine set in E^n
 - If $\text{rank}(A) = m$,
 - then $\dim(K) = n - m$



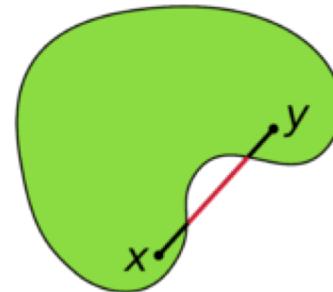
Convex set

- A set $C \subseteq E^n$ is convex if for any $x, y \in C$, and any $\lambda \in (0,1)$, we have

$$\lambda x + (1-\lambda)y \in C,$$



Convex



Not convex

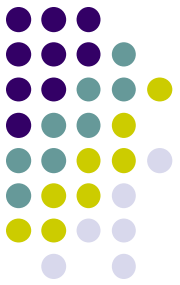
- A polyhedron is equal to the intersection of a finite number of closed halfspaces of E^n , indicated as $\{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$
- Feasible region $D = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$ is a convex set



Convex set and face

- If non-empty subset C' of convex set C satisfies the following conditions:
 - If $\mathbf{u} \in C'$, and \mathbf{u} is a point on a line in C , there exists $\mathbf{x}, \mathbf{y} \in C$ to make $\mathbf{u} \in (\mathbf{x}, \mathbf{y})$,
 - We can infer that $\mathbf{x}, \mathbf{y} \in C'$, C' is called a face of C .
 - Apparently, C is also a face.

Face, extreme point and extreme direction

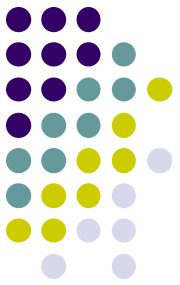


- Theorem: If and only if there exists an index set $Q \subseteq \{1, \dots, n\}$ to make

$$D' = \{\mathbf{x} | A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0, \text{ when } x_j = 0 \text{ and } j \in Q\}$$

Then D' is a subset and face of feasible region

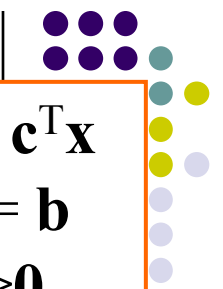
- If $\dim(D) = n - m$, D' is a face of D , and $\dim(D') = n - m - k$, then $|Q| \geq k$.
- 1-dimensional face of convex C is called edge, 0-dimensional face is called extreme point. If two extreme points are on the same line, they are said to be adjacent.
- Equivalent definition of extreme point: A vector $\mathbf{x} \in C$ is an extreme point, if $\mathbf{x} = \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2$, wherein $\mathbf{x}_1, \mathbf{x}_2 \in C$ and $\lambda \in (0, 1)$, there must be $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}$.
- Intuitively, extreme point is not an interior point on any hyperplane of C .



Extreme direction

- Let set $C \in E^n$ be convex set and $C \neq \emptyset$, $\mathbf{d} \in E^n$, $\mathbf{d} \neq \mathbf{0}$. If for any $\mathbf{x} \in C$, $\lambda > 0$, $\mathbf{x} + \lambda \mathbf{d} \in C$, \mathbf{d} is a extreme direction of C .
- If a direction \mathbf{d} can't be expressed as linear combination of two directions, \mathbf{d} is called extreme direction of C .
- The proposition that \mathbf{d} is the direction of $D = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\} \neq \emptyset$ is equivalent to:
 - \mathbf{d} satisfies conditions $A\mathbf{d} = \mathbf{0}$, $\mathbf{d} \geq \mathbf{0}$, $\mathbf{d} \neq \mathbf{0}$

Basic Feasible Solution


$$\begin{aligned}\min z &= \mathbf{c}^T \mathbf{x} \\ \text{s.t. } A\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0}\end{aligned}$$

Assume that $\text{rank}(A)=m$, then there must be m linearly independent column vectors constituting the full rank matrix B , the remaining columns of A constitute sub-matrix N , so $A=(B, N)$. Accordingly, $\mathbf{x}=(\mathbf{x}_B, \mathbf{x}_N)$, then $A\mathbf{x}=\mathbf{b}$ can be rewritten as

$$A\mathbf{x} = B\mathbf{x}_B + N\mathbf{x}_N = \mathbf{b}.$$

Because B is a full rank matrix, then there exists B^{-1} , so

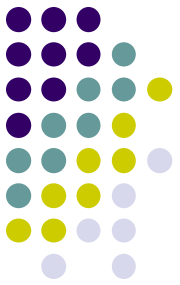
$$\mathbf{x}_B = B^{-1}\mathbf{b} - B^{-1}N\mathbf{x}_N$$

Choose \mathbf{x}_N arbitrarily, then obtain the corresponding \mathbf{x}_B ,

$$(\mathbf{x}_B, \mathbf{x}_N)$$

is a solution to $A\mathbf{x}=\mathbf{b}$.

Let $\mathbf{x}_N=0$, then $\mathbf{x}_B=B^{-1}\mathbf{b}$, $\mathbf{x}=(B^{-1}\mathbf{b}, 0)$ is called a basic solution to the constraint equations.

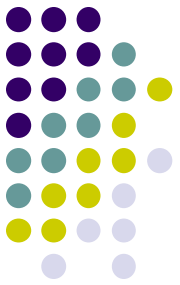


Basic feasible solution

- Suppose that $\text{rank}(A)=m$
 - m linearly independent columns of A constitute full rank matrix B
 - The remaining columns of A constitute sub-matrix N
 - $A=(B, N)$
- $\mathbf{x}=(\mathbf{x}_B, \mathbf{x}_N)$
 - $A\mathbf{x}=\mathbf{b}$ can be rewritten as $B\mathbf{x}_B + N\mathbf{x}_N=\mathbf{b}$

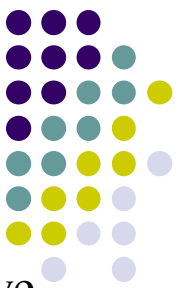
B is named basis, m linearly independent columns of B are named basic vectors, m components of \mathbf{x}_B are named basic variables, the remaining variables are called nonbasic variable.

A basic solution maybe not be nonnegative, so it may not feasible. If a basic solution is nonnegative, then it is a basic feasible solution, and name B feasible basis.



Basic feasible solution

- Theorem:
 - If the feasible solution \mathbf{x} is a basic feasible solution, the columns corresponding to positive components of \mathbf{x} are linearly independent.
 - If the feasible solution \mathbf{x} is a basic feasible solution, \mathbf{x} is an extreme point of D .



Proof: (sufficiency)

Let x be an extreme point of D , top k components be positive, we are aim to prove that a_1, \dots, a_k are linearly independent.

Reduction to absurdity. If they are linearly dependent, there exists nonzero $y = (y_1, \dots, y_k, 0, \dots, 0)^T$, to make

$$\sum_{j=1}^k y_j a_j = 0.$$

For any positive real number δ , we have

$$\sum_{j=1}^k (x_j \pm \delta y_j) a_j = b$$

Then take $x^1 = x + \delta y$, $x^2 = x - \delta y$,

It's easy to verify that when δ is small enough, $x^1, x^2 \in D$.

Because $y \neq 0$, then $x^1 \neq x^2$, while

$$x = \frac{1}{2} x^1 + \frac{1}{2} x^2, \text{ this conflicts with that } x \text{ is an extreme point.}$$

Therefore, a_1, \dots, a_k are linearly independent, x is a basic feasible solution.



- Proof: (necessity)

Let x be a basic feasible solution, top k components be positive.

Suppose that there exists $x^1, x^2 \in D$ and $\lambda \in (0,1)$, to make

$$x = \lambda x^1 + (1 - \lambda) x^2$$

Because when $j \geq k + 1$, $x_j = 0$, there is $x_j^1 = x_j^2 = 0$. According to $Ax^1 = Ax^2 = b$, we obtain that,

$$\sum_{j=1}^k (x_j^1 - x_j^2) a_j = 0$$

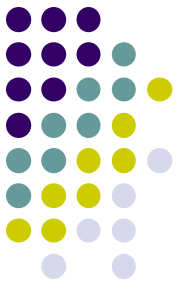
Because x is a basic feasible solution, and a_1, \dots, a_k are linearly independent. Therefore, $x_j^1 = x_j^2$ ($j = 1, \dots, k$), so as to $x^1 = x^2 = x$.
 x is an extreme point.



Degeneracy

- If the basic feasible solution \mathbf{x} , all its basic variables are positive, then \mathbf{x} is non-degenerated; Otherwise, it is degenerated.
- A feasible basis uniquely determines a basic feasible solution; Otherwise, if a basic feasible solution is non-degenerated, it uniquely corresponds to a feasible basis.
- If a basic feasible solution is degenerated, the corresponding basis is not unique.
- If all the basic feasible solutions to a LP problem are non-degenerated, then the problem is non-degenerated, otherwise it is degenerated.

Extreme direction of feasible region



- Algebraic properties of extreme direction: The direction \mathbf{d} of $D = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ has k nonzero components, then the sufficient and necessary condition that \mathbf{d} is the extreme direction of D is that the rank of the columns corresponding to nonzero components in \mathbf{d} is $k-1$
- Geometric properties of extreme direction: The sufficient and necessary condition that \mathbf{d} is an extreme direction of D is that \mathbf{d} is a direction of a certain half-line face in D .
- If D has extreme directions, D is a boundless set; Conversely, if D is a boundless set, D has directions and they are extreme directions.

Linear programming basic theorem



Basic theorem 1 (Representation theorem)

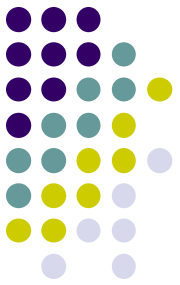
Theorem: Let all extreme points of $D = \{x \mid Ax = b, x \geq 0\}$ be x^1, \dots, x^k , all extreme directions be d^1, \dots, d^l , the sufficient and necessary condition to $x \in D$ is that there exists a set of $\lambda_i (i = 1, \dots, k)$ and $\mu_j (j = 1, \dots, l)$ to make

Extreme point **Extreme direction**

$$x = \sum_{i=1}^k \lambda_i x^i + \sum_{j=1}^l \mu_j d^j,$$

$$\lambda_i \geq 0, i = 1, \dots, k; \mu_j \geq 0, j = 1, \dots, l$$

$$\sum_{i=1}^k \lambda_i = 1$$



Basic Theorem 2

Given a LP problem

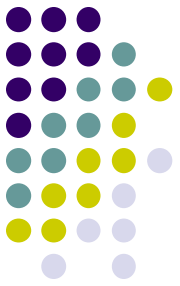
$$\min z = \mathbf{c}^T \mathbf{x}$$

$$\text{s.t. } A\mathbf{x} = \mathbf{b}$$

$$\mathbf{x} \geq \mathbf{0}$$

(1) If there exists a finite optimal solution, the optimal solution must be an extreme point in the feasible region D .

(2) The sufficient and necessary condition that objective function has optimal solution is that for any extreme direction \mathbf{d}^j of D we have $\mathbf{c}\mathbf{d}^j \geq 0$.



Proof: Let all extreme points in $D = \{x \mid Ax = b, x \geq 0\}$ be x^1, \dots, x^k , all extreme directions be d^1, \dots, d^l , according to representation theorem, we know that any $x \in D$ can be expressed as

$$x = \sum_{i=1}^k \lambda_i x^i + \sum_{j=1}^l \mu_j d^j,$$
$$\lambda_i \geq 0, i = 1, \dots, k; \quad \mu_j \geq 0, j = 1, \dots, l$$
$$\sum_{i=1}^k \lambda_i = 1$$

So transform the *LP* problem about x into the one about λ and μ .

$$\min \sum_{i=1}^k (c^T x^i) \lambda_i + \sum_{j=1}^l (c^T d^j) \mu_j$$
$$\text{s.t. } \sum_{i=1}^k \lambda_i = 1; \quad \lambda_i \geq 0, i = 1, \dots, k; \quad \mu_j \geq 0, j = 1, \dots, l$$

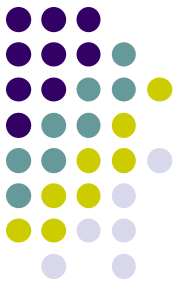
Because μ_j can be arbitrarily large, to realize the objective function has lower boundary, there must be

$$c^T d^j \geq 0 \quad (j = 1, \dots, l)$$

or there are no extreme direction.

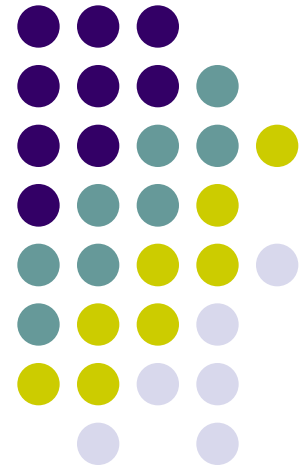
Let $c^T x^p = \min \{c^T x^i \mid 1 \leq i \leq k\}$, then we know that the objective function is not less than $c^T x^p$. So let $\lambda_p = 1$ the remaining $\lambda_i = 0$, that is the objective function reaches the minimum. That is to say we have verified that the function reaches the optimal at the extreme point x_p .

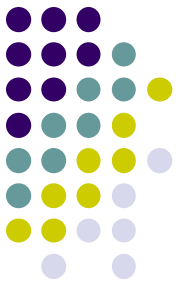
Annotation and Interpretation of The Theorem



- Explanation of the theorem:
 - The optimal to the objective function of stand form LP problem must be at a basic feasible solution.
 - To solve stand form LP problem, only need to search in the set of basic feasible solutions.
- The way that work out and compare all the basic feasible solutions is not feasible. And when n is large, the number of basic feasible solution is quite large.
- A general method is called the somplex method, in which we only need to search in a subset of basic feasible solutions.

Simplex Method





Basic simplex method

- Main idea: First, find a basic feasible solution, and discriminate whether it is optimal. If not, find a better basic feasible solution and discriminate it again. The method continues until finding the optimal solution or discriminating it is boundless.
- Two major problems:
 - Find initial solution
 - The way to discrimination and iteration (consider first)



Discrimination and iteration

- Suppose that $\text{rank}(A) = m < n$, and we have found a non-degenerated feasible basic solution, i.e. found a basis B . Then rewrite $A\mathbf{x}=\mathbf{b}$ as

$$\mathbf{x}_B + B^{-1}N\mathbf{x}_N = B^{-1}\mathbf{b} \quad (1)$$

Denote $B = (a_1, \dots, a_m)$

$$\bar{\mathbf{a}}_j = B^{-1}\mathbf{a}_j = (\bar{a}_{1j}, \dots, \bar{a}_{mj})^T, j = 1, \dots, n$$

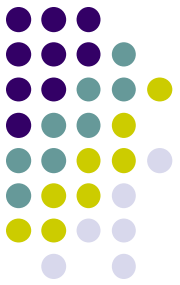
$$\bar{\mathbf{b}} = B^{-1}\mathbf{b} = (\bar{b}_1, \dots, \bar{b}_m)^T$$

$$\bar{N} = B^{-1}N$$

$$\text{Then rewrite formula (1) as } \mathbf{x}_B + \bar{N}\mathbf{x}_N = \bar{\mathbf{b}} \quad (2)$$

Obviously, different bases correspond to different equations.

Formula (2) is called canonical system of equations, canonical form for short.



Discrimination and iteration

$$\mathbf{x}_B + \bar{N}\mathbf{x}_N = \bar{\mathbf{b}} \quad (2-2)$$

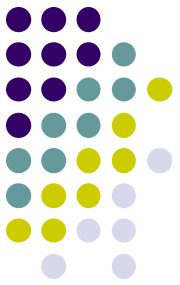
Obviously, different bases correspond to different equations. Formula (2) is called canonical system of equations of basis B, canonical form for short.

If $\bar{\mathbf{b}} \geq 0$, then formula (2) corresponds to the feasible basic solution $\bar{\mathbf{x}} = \begin{pmatrix} \bar{\mathbf{b}} \\ 0 \end{pmatrix}$. Transform the objective function,

$$z = \mathbf{c}\mathbf{x} = \mathbf{c}_B\bar{\mathbf{b}} - (\mathbf{c}_B\bar{N} - \mathbf{c}_N)\mathbf{x}_N = \mathbf{c}_B\bar{\mathbf{b}} - \sum_{j=m+1}^n (\mathbf{c}_B\bar{\mathbf{a}}_j - c_j)\mathbf{x}_j \quad (3)$$

Substitute z_0 for the objective function value at $\bar{\mathbf{x}}$,

then $z_0 = \mathbf{c}\bar{\mathbf{x}} = \mathbf{c}_B\bar{\mathbf{b}}$ (i.e. the constant term of the equation above)



Discrimination and iteration

$\bar{\mathbf{a}}_1, \dots, \bar{\mathbf{a}}_m$ are linearly independent,

so when $j = 1, \dots, m$

$$\mathbf{c}_B \bar{\mathbf{a}}_j - c_j = 0.$$

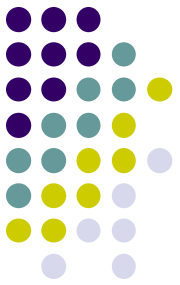
Introduce the notation

$$\zeta_j = \mathbf{c}_B \bar{\mathbf{a}}_j - c_j, j = 1, \dots, n$$

or in the form of vectors

$$\zeta = \mathbf{c}_B B^{-1} A - \mathbf{c} = (\zeta_B, \zeta_N) = (\mathbf{0}, \mathbf{c}_B B^{-1} N - \mathbf{c}_N)$$

Then rewrite equation (3) as $z = \mathbf{c}_B \bar{\mathbf{b}} - \zeta \mathbf{x}$.



Discrimination and iteration

$$z = \mathbf{c}_B \bar{\mathbf{b}} - \zeta \mathbf{x}, \quad \zeta = \mathbf{c}_B B^{-1} A - \mathbf{c} = (\zeta_B, \zeta_N) = (\mathbf{0}, \mathbf{c}_B B^{-1} N - \mathbf{c}_N)$$

After variable transformation, describe *LP* problem as

$$\min z = z_0 - \zeta \mathbf{x} \quad (4)$$

$$s.t. \quad \mathbf{x}_B + B^{-1} N \mathbf{x}_N = \bar{\mathbf{b}}, \quad (5)$$

$$\mathbf{x} \geq \mathbf{0}.$$

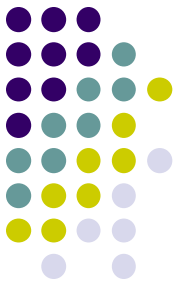
Theorem 2.1: If ζ in formula(4) satisfies $\zeta \leq 0$, $\bar{\mathbf{x}}$ is the optimal.

Theorem 2.2: If there exists a component in ζ $\zeta_k > 0$, and $\bar{\mathbf{a}}_k \leq \mathbf{0}$, then there is no solution to the primal problem.

Theorem 2.3: If there is $\zeta_k > 0$ in formula(4), and there is at least one positive component in $\bar{\mathbf{a}}_k$,

then we can find the basic feasible solution $\hat{\mathbf{x}}$ to make $\hat{\mathbf{x}} < \mathbf{c} \bar{\mathbf{x}}$.

Discrimination and iteration



Construct(a better basic feasible solution) $\hat{\mathbf{x}}$. Let

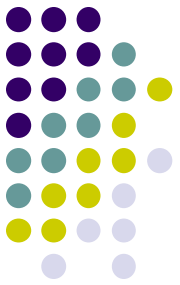
$$\theta = \min \left\{ \frac{\bar{b}_i}{\bar{a}_{ik}} \mid \bar{a}_{ik} > 0, i = 1, \dots, m \right\} = \frac{\bar{b}_r}{\bar{a}_{rk}} > 0$$

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \theta \left(\begin{pmatrix} -\bar{\mathbf{a}}_k \\ 0 \end{pmatrix} + \mathbf{e}_k \right) = \begin{pmatrix} \bar{\mathbf{b}} - \theta \bar{\mathbf{a}}_k \\ 0 \end{pmatrix} + \theta \mathbf{e}_k.$$

The following will prove $\hat{\mathbf{x}}$ is a feasible solution. Plug in the formula(5),

$$\hat{\mathbf{x}}_B + N\hat{\mathbf{x}}_N = \bar{\mathbf{b}} - \theta \bar{\mathbf{a}}_k + \sum_{j=m+1}^n \hat{x}_j \bar{\mathbf{a}}_j = \bar{\mathbf{b}} - \theta \bar{\mathbf{a}}_k + \theta \bar{\mathbf{a}}_k = \bar{\mathbf{b}}. (\text{Notice } \hat{x}_N = (0, \dots, \theta, 0 \dots 0))$$

Due to θ construction, it's clear that $\hat{\mathbf{x}} \geq 0$. So $\hat{\mathbf{x}}$ is a feasible solution.



Discrimination and iteration

The following will prove $\hat{\mathbf{x}}$ is a basic solution. Each component of $\hat{\mathbf{x}}$ is

$$\hat{x}_i = \bar{b}_i - \left(\bar{b}_r / \bar{a}_{rk} \right) \bar{a}_{ik}, i = 1, \dots, m$$

$$\hat{x}_k = \bar{b}_k / \bar{a}_{rk}$$

$$\hat{x}_j = 0, j = m + 1, \dots, n, j \neq k.$$

Because of $\hat{x}_r = 0$, while $\hat{x}_k > 0$,

So only need to prove $a_1, \dots, a_{r-1}, a_k, a_{r+1}, \dots, a_m$ are linearly independent.



Proof: $a_1, \dots, a_{r-1}, a_k, a_{r+1}, \dots, a_m$ are linearly independent

Proof by contradiction. If they are linearly dependent, because of $\{\mathbf{a}_1, \dots, \mathbf{a}_m\}$ linearly dependent, \mathbf{a}_k must be a linear combination of the other $m-1$ vectors.

$$\text{i.e. } \mathbf{a}_k = \sum_{i=1, i \neq r}^m y_i \mathbf{a}_i$$

$$\text{Additionally, } \bar{\mathbf{a}}_k = B^{-1} \mathbf{a}_k, \text{ so as to } \mathbf{a}_k = B \bar{\mathbf{a}}_k = \sum_{i=1}^m \bar{a}_{ik} \mathbf{a}_i$$

Two equations above subtracting to yield

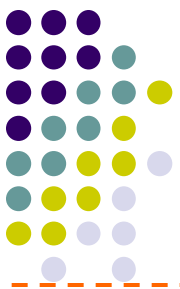
$$\bar{a}_{rk} \mathbf{a}_r + \sum_{i=1, i \neq r}^m (\bar{a}_{ik} - y_i) \mathbf{a}_i = \mathbf{0}$$

Linearly dependent

Because of $\bar{a}_{rk} \neq 0$, $\mathbf{a}_1, \dots, \mathbf{a}_m$ are linearly independent, they're contradictory.

Finally,

$$z(\hat{\mathbf{x}}) = z_0 - \zeta \hat{\mathbf{x}} = z_0 - \zeta_k \theta < z_0.$$

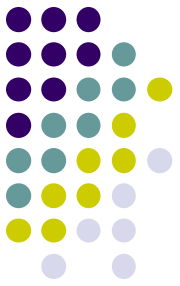


Simplex method procedure

ζ is called test vector. In iteration process, if ζ has more than one positive component, then choose column \mathbf{a}_k corresponding to the maximum component ζ_k to enter the basis, making objective function descend more quickly.

1. Find an initial feasible basis;
2. Compute the corresponding canonical form;
3. Compute $\zeta_k = \max \{\zeta_j \mid j = 1, \dots, n\}$
4. If $\zeta_k \leq 0$, terminate. The optimum solution $\mathbf{x} = \begin{pmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{b}} \\ \mathbf{0} \end{pmatrix}$
and optimum value $z = \mathbf{c}_B \bar{\mathbf{b}}$ have been found.
5. If $\bar{\mathbf{a}}_k \leq \mathbf{0}$, terminate, primal problem is boundless.
6. Compute the minimum ratio $\frac{\bar{b}_r}{\bar{a}_{rk}} = \min \left\{ \frac{\bar{b}_i}{\bar{a}_{ik}} \mid \bar{a}_{ik} > 0 \right\}$
7. Replace \mathbf{a}_{Br} with \mathbf{a}_k to obtain a new basis, go to step 2.

Simplex tableau



According to the way to construct θ and \hat{x} , we obtain the modified

basic feasible solution, i.e. transform the nonbasic variable x_k into positive basic variable.

Meanwhile, make the basic variable x_r zero (turn into nonbasic variable).

Equivalently, basis $B = [a_1, \dots, a_m]$ turns into another basis

$$\hat{B} = [a_1, \dots, a_r, a_k, a_{r+1}, \dots, a_m].$$

*Basis transform

* x_k is called entering basis variable, x_r is exiting basis variable.

With basis B , the coefficient of canonical form augmented matrix is,

x_1	...	x_r	...	x_m	x_{m+1}	...	x_k	...	x_n	
1	...	0	...	0	$\bar{a}_{1,m+1}$...	$\bar{a}_{1,k}$...	$\bar{a}_{1,n}$	\bar{b}_1
\vdots	\ddots	\vdots		\vdots	\vdots		\vdots		\vdots	\vdots
0	...	1	...	0	$\bar{a}_{r,m+1}$...	$\bar{a}_{r,k}$...	$\bar{a}_{r,n}$	\bar{b}_r
\vdots		\vdots	\ddots	\vdots	\vdots		\vdots		\vdots	\vdots
0	...	0	...	1	$\bar{a}_{m,m+1}$...	$\bar{a}_{m,k}$...	$\bar{a}_{m,n}$	\bar{b}_m

Exchange x_r and x_k .

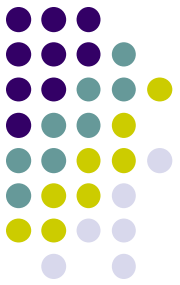
Transform the k th column into unit vector e_r by elementary transformation.

1) Divide the r th row \bar{a}_r^T by \bar{a}_{rk} , $\hat{a}_r^T = \bar{a}_r^T / \bar{a}_{rk}$

2) Subtract the new r th row \hat{a}_r^T multiplying with \bar{a}_{ik} from the i th row ($i \neq r$) \bar{a}_i^T , $\hat{a}_i^T = \bar{a}_i^T - \hat{a}_r^T \bar{a}_{ik}$

The last column is the basic feasible solution corresponding to basis \hat{B}

$$\hat{b}_r = \bar{b}_r / \bar{a}_{rk}, \hat{b}_i = \bar{b}_i - (\bar{b}_r / \bar{a}_{rk}) \bar{a}_{ik}, i \neq r$$



When changing basis, change B into \hat{B} in the objective function $z = c_B \bar{b} - (c_B \bar{N} - c_N) x_N$.

Regard the following equations (equivalent form of the above equation)

$$z + (c_B B^{-1} N - c_N) x_N = c_B \bar{b}, \text{ or } z + \zeta_N x_N = z_0$$

as an equation, make elementary transformation both to this equation and canonical form.

	z	x_1	\dots	x_r	\dots	x_m	x_{m+1}	\dots	x_k	\dots	x_n	RHS
z	1	0	\dots	0	\dots	0	ζ_{m+1}	\dots	ζ_k	\dots	ζ_n	z_0
x_1	0	1	\dots	0	\dots	0	$\bar{a}_{1,m+1}$	\dots	$\bar{a}_{1,k}$	\dots	$\bar{a}_{1,n}$	\bar{b}_1
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
x_r	0	0	\dots	1	\dots	0	$\bar{a}_{r,m+1}$	\dots	$\bar{a}_{r,k}$	\dots	$\bar{a}_{r,n}$	\bar{b}_r
\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
x_m	0	0	\dots	0	\dots	1	$\bar{a}_{m,m+1}$	\dots	$\bar{a}_{m,k}$	\dots	$\bar{a}_{m,n}$	\bar{b}_m

Simplex tableau

Pivot column

Pivot row

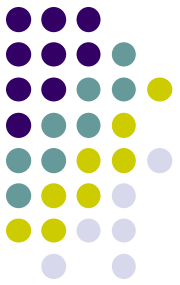
Pivot element

Denote simplex tableau as

	z	x_B	x_N	RHS
z	1	0	ζ_N	z_0
x_B	0	I	\bar{N}	\bar{b}

The above transformation is called pivot, similar to principal component elimination method to solve linear equations.

Initial solution: Two-phase simplex method



Let the original problem be

$$\min \mathbf{c}\mathbf{x}$$

$$s.t. \quad A\mathbf{x} = \mathbf{b} \quad (\mathbf{b} \geq \mathbf{0})$$

$$\mathbf{x} \geq \mathbf{0}$$

Introduce m artificial variables $\mathbf{x}_a = (x_{n+1}, \dots, x_{n+m})^T$,

consider the following auxiliary problem

$$\min g = \mathbf{1}\mathbf{x}_a$$

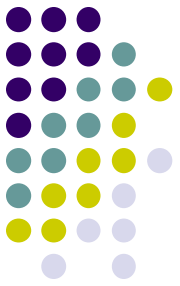
$$s.t. \quad A\mathbf{x} + \mathbf{x}_a = \mathbf{b}$$

$$\mathbf{x}, \mathbf{x}_a \geq \mathbf{0}$$

wherein $\mathbf{1} = (1, \dots, 1)$. Let D and D' respectively be the feasible region of the primal problem and the auxiliary problem,

Obviously, $\mathbf{x} \in D \Leftrightarrow \begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix} \in D'$. And $\begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix} \in D' \Leftrightarrow \min g = 0$.

Initial solution: Two-phase simplex method



m artificial variables are $\mathbf{x}_a = (x_{n+1}, \dots, x_{n+m})^T$, consider the following auxiliary problem:

$$\min g = \mathbf{1}\mathbf{x}_a$$

$$s.t. \quad A\mathbf{x} + \mathbf{x}_a = \mathbf{b}$$

$$\mathbf{x}, \mathbf{x}_a \geq 0$$

$$\begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix} \in D' \Leftrightarrow \min g = 0.$$

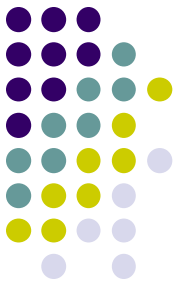
$$\begin{pmatrix} \mathbf{x} \\ \mathbf{x}_a \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{b} \end{pmatrix} \text{ is a basic feasible solution to the auxiliary problem,}$$

Then we can solve the problem iteratively by simplex method, there are two possible results :

1) $\min g > 0$. Note that $D = \emptyset$.

2) $\min g = 0$. It's quite natural that $\mathbf{x}_a = 0$, we get a feasible solution to the primal problem after removing it.

Initial Solution: Two-phase simplex method



m artificial variables $\mathbf{x}_a = (x_{n+1}, \dots, x_{n+m})^T$, we obtain the auxiliary problem

$$\min g = \mathbf{1}\mathbf{x}_a$$

$$s.t. \quad A\mathbf{x} + \mathbf{x}_a = \mathbf{b}$$

$$\mathbf{x}, \mathbf{x}_a \geq 0$$

$$\begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix} \in D' \Leftrightarrow \min g = 0.$$

- * If all the artificial variables are nonbasic variables, the solution is a basic feasible solution.
- * Otherwise, eliminate artificial variables from basic variables by pivot.

Let basic variable x_r be an artificial variable,

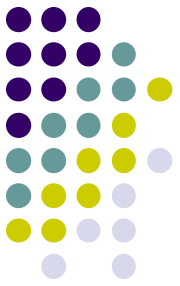
▷ Transform with top n nonzero elements in the r th row as pivot element, then

non-artificial variable x_s enters the basis as well as

eliminate x_r from basic variables.

▷ If top n elements in the r th row are zero, eliminate this row and its corresponding artificial variables directly.

Initial Solution: The Big-M method



Let the primal problem be

$$\min cx$$

$$s.t. \ Ax = b \ (b \geq 0)$$

$$x \geq 0$$

Introduce $x_a \in E^m$ and a large positive constant M , we obtain the new problem as follows:

$$\min cx + M1x_a$$

$$s.t. \ Ax + x_a = b$$

$$x, x_a \geq 0$$

Wherein $1 = (1, \dots, 1)$. When M is large enough,

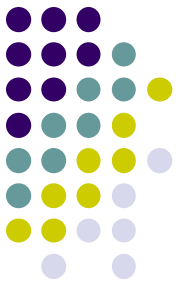
x is the optimal solution to the primal problem $\Leftrightarrow \begin{pmatrix} x \\ 0 \end{pmatrix}$ is the optimal solution to the new problem.

And the new problem has initial solution $\begin{pmatrix} 0 \\ b \end{pmatrix}$, so it can be solved by the simplex method.

Degeneracy and anticycling rule



- If there exists zero component among the vector right to simplex tableau, i.e. the solution is degenerate, the algorithm may loop indefinitely. We add some pivoting rules to avoid cycling.
 - Dictionary ordered method
 - Bland criteria



Dictionary ordered method

Nonzero vector x , and its first nonzero component is nonnegative, called dictionary sequence nonnegative, denoted as $x \succcurlyeq 0$.

Vectors x, y , if $x - y \succcurlyeq 0$, x is greater than or equal to y according to dictionary sequence.

If a set of x^i , there exists a x^t , for all x^i satisfy $x^i \succcurlyeq x^t$, then say x^t is the minimum one among these vectors according to dictionary sequence.

Denoted as $x^t = \text{lex min } x^i$.

The so-called dictionary ordered method is as follows:

After the choice of entering variable x_k , let

$$\frac{p_r}{\bar{a}_{rk}} = \text{lex min} \left\{ \frac{p_i}{\bar{a}_{ik}} \mid \bar{a}_{ik} > 0, i = 1, \dots, m \right\},$$

wherein p_i is the i th row of (\bar{b}, B^{-1}) .

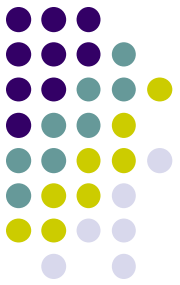


Bland criteria

- 1) Choose nonbasic variable x_k , which corresponds to the positive test number ζ_k with the minimum subscript, as entering basis variable.
- 2) Determine exiting variable x_l : If there are several $\frac{b_r}{\bar{a}_{rk}}$ reaching minimum, choose the basic variable with the minimum subscript as exiting basis variable. i.e.

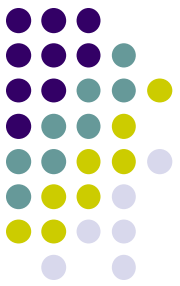
$$l = \min \left\{ r \mid \frac{b_r}{\bar{a}_{rk}} = \min \left\{ \frac{b_i}{\bar{a}_{ik}} \mid \bar{a}_{ik} > 0 \right\} \right\}$$

- The advantage of Bland rules is simple and practicable, but it only cares about the minimal subscript, without considering the decline speed of objective function. So its efficiency is lower than dictionary ordered method and simplex method.
- In practice, degeneration is usual, but degeneration doesn't bring about loop necessarily. In fact, loop phenomenon is unusual.



Revised simplex method

- When a LP problem is quite large in dimension, we must consider how to reduce the storage and computational burden. In practice, we usually use revised simplex method.



Inverse matrix method

General simplex tableau is as follows :

$$\begin{bmatrix} & z & \mathbf{x}_B & \mathbf{x}_N & RHS \\ z & 1 & 0 & \mathbf{c}_B B^{-1} \mathbf{N} - \mathbf{c}_N & \mathbf{c}_B B^{-1} \mathbf{b} \\ \mathbf{x}_B & 0 & I & B^{-1} \mathbf{N} & B^{-1} \mathbf{b} \end{bmatrix}$$

Inverse matrix method only store these data as follows in each simplex tableau:

$$\begin{bmatrix} w & z_o \\ B^{-1} & \bar{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} \mathbf{c}_B B^{-1} & \mathbf{c}_B B^{-1} \mathbf{b} \\ B^{-1} & B^{-1} \mathbf{b} \end{bmatrix}, \text{ which is called revised simplex tableau.}$$

We compute the original data iteratively according to this tableau. Determine Take x_k as entering variable according to $\zeta_k = \max \{ w a_j - c_j \mid x_j \text{ nonbasic variable} \}$

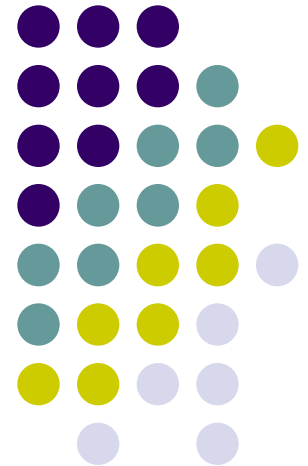
Compute $\bar{\mathbf{a}}_k = B^{-1} \mathbf{a}_k$, determine the exiting variable x_r by the smallest ratio, then obtain new basis \hat{B} , finally construct the revised simplex tableau corresponding to \hat{B} .

If we must construct B^{-1} at each iteration, the calculation is very large.

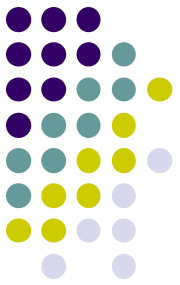
Theorem: add a column $\begin{pmatrix} \zeta_k \\ \bar{\mathbf{a}}_k \end{pmatrix}$ right to the revised simplex tableau corresponding to B , pivot

with \bar{a}_{rk} as pivot element, then we obtain the revised simplex tableau corresponding to basis \hat{B} .

Optimality Conditions and Duality Theory

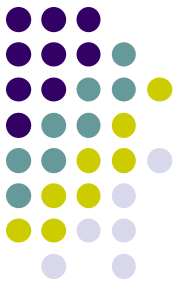


Duality



- Given a linear programming problem P , we can associate it with another (dual) linear programming D . People can solve the primal problem P by solving the dual problem D .

Karush-Kuhn-Tucker condition



- Theorem

Let A a $m \times n$ matrix, $b \in E^m$, $c \in E^n$, $x \in E^n$, LP problem:

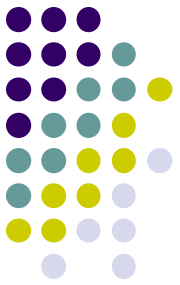
$$\begin{aligned} \min & c^T x \\ \text{s.t.} & Ax \geq b \\ & x \geq 0 \end{aligned}$$

x^* is the optimal solution, which its sufficient and necessary condition is that there exists $w \in E^m$, $v \in E^n$ making the following *KKT* conditions be satisfied

$$\begin{aligned} Ax^* &\geq b, x^* \geq 0 \\ c - w^T A - v &= 0, w \geq 0, v \geq 0 \\ w^T (Ax^* - b) &= 0, v^T x^* = 0 \end{aligned}$$

If v is omitted (by the equation $c - w^T A - v = 0$), we obtain another form of *KTT* condition

$$\begin{aligned} Ax^* &\geq b, x^* \geq 0 \\ wA &\leq c, w \geq 0 \\ w(Ax^* - b) &= 0, (wA - c)x^* = 0 \end{aligned}$$



Kuhn Tucker condition

The $K - T$ condition for the standard form LP problem

As long as rewrite $Ax = b$ as $\begin{pmatrix} A \\ -A \end{pmatrix} x \geq \begin{pmatrix} b \\ -b \end{pmatrix}$, then we

can infer the corresponding $K - T$ condition

$$Ax = b, x \geq 0$$

$$c - wA - v = 0, v \geq 0$$

$$vx = 0$$

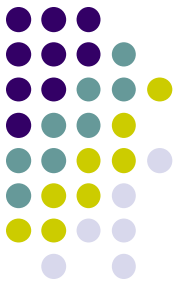
or the equivalent form

$$Ax = b, x \geq 0$$

$$wA \leq c$$

$$(wA - c)x = 0.$$

Kuhn Tucker condition



The sufficient and necessary condition that for a feasible basic solution $\begin{pmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{b}} \\ 0 \end{pmatrix} (\bar{\mathbf{b}} > 0)$

is optimal is that the test vector $\zeta \leq 0$, and what's the relation with $K-T$ condition?

Because of $\mathbf{v}\mathbf{x} = \mathbf{v}_B\mathbf{x}_B + \mathbf{v}_N\mathbf{x}_N = 0$, and $\mathbf{x}_B > 0$, we infer that $\mathbf{v}_B = 0$.

Condition $\mathbf{c} - \mathbf{w}\mathbf{A} - \mathbf{v} = 0$ can be rewritten as $(\mathbf{c}_B, \mathbf{c}_N) - \mathbf{w}(B, N) - (\mathbf{v}_B, \mathbf{v}_N) = (0, 0)$, so as to

$$\begin{cases} \mathbf{c}_B - \mathbf{w}B = 0 \\ \mathbf{c}_N - \mathbf{w}N - \mathbf{v}_N = 0 \end{cases} \Rightarrow \begin{cases} \mathbf{w} = \mathbf{c}_B B^{-1} \\ \mathbf{v}_N = \mathbf{c}_N - \mathbf{w}N = \mathbf{c}_N - \mathbf{c}_B B^{-1}N \end{cases}$$

i.e. given a basic feasible solution \mathbf{x} , if we let

$$\mathbf{v}_B = 0, \mathbf{w} = \mathbf{c}_B B^{-1}, \mathbf{v}_N = \mathbf{c}_N - \mathbf{w}N = \mathbf{c}_N - \mathbf{c}_B B^{-1}N,$$

except $\mathbf{v} \geq 0$, the rest $K-T$ conditions are satisfied.

And according to the choice of \mathbf{v} , it's easy to know $\mathbf{v} \geq 0 \Leftrightarrow \zeta \leq 0$.



Duality theorem

- Regard the dual problem as transposition of the primal problem

$$\min \mathbf{c}^T \mathbf{x} \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0$$

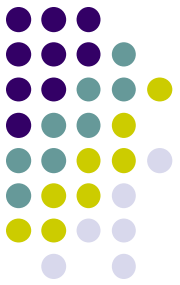
$$\min c^T x \quad \text{s.t.} \quad \begin{bmatrix} A \\ -A \end{bmatrix} x \geq \begin{bmatrix} b \\ -b \end{bmatrix}, x \geq 0$$

- The dual problem is:

$$\max \begin{bmatrix} b^T & -b^T \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad \text{s.t.} \quad \begin{bmatrix} A^T & -A^T \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \leq c$$

- Wherein y_1 and y_2 are dual variables corresponding to constraints $\mathbf{A}\mathbf{x} \geq \mathbf{b}$ and $-\mathbf{A}\mathbf{x} \leq \mathbf{B}$, respectively. Let $\mathbf{y} = \mathbf{y}_1 - \mathbf{y}_2$

$$\max \mathbf{b}^T \mathbf{y} \quad \text{s.t.} \quad \mathbf{A}^T \mathbf{y} \leq \mathbf{c}$$



Duality theorem

- Let \mathbf{x} and \mathbf{w} be the feasible solution to (P) and (D), respectively. So as to $\mathbf{c}\mathbf{x} \geq \mathbf{w}\mathbf{b}$.
- Let \mathbf{x}^* and \mathbf{w}^* be the feasible solution to (P) and (D), respectively. So as to the sufficient and necessary condition is $\mathbf{c}\mathbf{x}^* = \mathbf{w}^*\mathbf{b}$, for which \mathbf{x}^* and \mathbf{w}^* are optimal to (P) and (D), respectively.
- Corollary: If (P) has an optimal solution \mathbf{x}^* , so does its dual (D), optimal solution \mathbf{w}^* , and $\mathbf{c}\mathbf{x}^* = \mathbf{w}^*\mathbf{b}$; If (P) is boundless, (D) has no solution and vice versa.
- Complementary slackness: Let \mathbf{x}^* and \mathbf{w}^* be optimal solution to the primal and the dual problem, respectively. So as to $\mathbf{w}^*(\mathbf{A}\mathbf{x}^* - \mathbf{b}) = 0$, $(\mathbf{c} - \mathbf{w}^*\mathbf{A})\mathbf{x}^* = 0$.

$K - T$ conditions:

$$\mathbf{A}\mathbf{x}^* \geq \mathbf{b}, \quad \mathbf{x}^* \geq 0$$

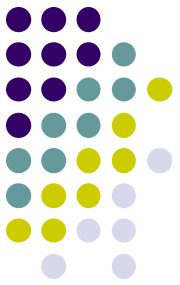
(primal feasibility)

$$\mathbf{w}\mathbf{A} \leq \mathbf{c}, \quad \mathbf{w} \geq 0$$

(dual feasibility)

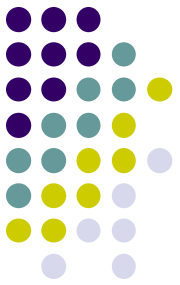
$$\mathbf{w}(\mathbf{A}\mathbf{x}^* - \mathbf{b}) = 0, \quad (\mathbf{w}\mathbf{A} - \mathbf{c})\mathbf{x}^* = 0$$

(complementary slackness)



Duality theorem

- The relation between K-T conditions and the optimal criterion of simplex condition: Simplex method satisfies all K-T condition except for condition $\mathbf{v} \geq \mathbf{0}$. i.e. simplex method keeps primal feasibility and complementary slackness, and modifies the feasible basic solution under the condition $\mathbf{c} - \mathbf{w}A - \mathbf{v} = \mathbf{0}$, to make $\mathbf{v} = \mathbf{c} - \mathbf{w}A \geq \mathbf{0}$ in dual feasibility be satisfied.
- Inspiration: When dual feasibility and complementary slackness are satisfied, we can modify the solution to satisfy primal feasibility.



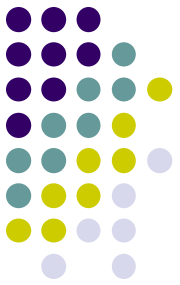
Dual simplex method

$$\begin{array}{ll} (P) \min cx & (D) \max wb \\ s.t. Ax = b & s.t. wA \leq c \\ x \geq 0 & \end{array}$$

The basic feasible solution to D : Let $A=(B, N)$, wherein B is a full rank matrix, then the solution $\bar{w} = c_B B^{-1}$ to $wB = c_B$ is a basic solution to (D) ,
If $\bar{w}N \leq c_N$, \bar{w} is the basic feasible solution to (D) .

The regular solution to P : If the test vector ζ corresponding to the basic solution $x = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$ to the primal (P) satisfies $\zeta = (0, c_B B^{-1}N - c_N) \leq 0$, x is called regular solution to problem (P) .

Basis B is called regular basis. It's easy to verify that the basic feasible solutions to D and the regular solutions to P are one-to-one relationship. The same as simplex method, to solve the dual programming (D) , the method moves from one basic feasible solution to another to increase the objective function value. Equivalently, to solve the primal programming (P) , the method moves one regular solution to another to increase the objective function value $z = wb = c_B B^{-1}b = cx$, until $B^{-1}b \geq 0$. i.e. when regular solution satisfies primal feasibility, the optimal solution is found. We call this method dual simplex method.



Dual simplex method

$$(P) \min \mathbf{c}\mathbf{x}$$

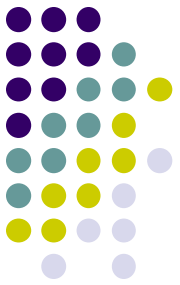
$$s.t. \mathbf{A}\mathbf{x} = \mathbf{b}$$

$$\mathbf{x} \geq 0$$

$$(D) \max \mathbf{w}\mathbf{b}$$

$$s.t. \mathbf{w}\mathbf{A} \leq \mathbf{c}$$

- 1) Find a normal basis B , construct its corresponding simplex tableau.
- 2) If $\bar{\mathbf{b}} = B^{-1}\mathbf{b} \geq 0$, terminate, the primal optimal solution has been found;
Otherwise, let $\bar{b}_r = \min \{\bar{b}_i \mid i = 1, \dots, m\}$.
- 3) If $\bar{a}^r \geq 0$, terminate, the primal has no optimal solution;
Otherwise, let $\frac{\zeta_k}{\bar{a}_{rk}} = \min \left\{ \frac{\zeta_j}{\bar{a}_{rj}} \mid a_{rj} < 0 \right\}$
- 4) Pivot with pivot element \bar{a}_{rk} , go to 2.



Primal-Dual simplex method

- Solve both the dual problem and the auxiliary problem at the first phase. Start with any feasible solution to the dual. In iterations, dual feasibility, complementary slackness and $x \geq 0$ are maintained. In this case, eliminate artificial variables to satisfy $Ax=b$.

$$\begin{array}{ll}
 (P) \min cx & (D) \max wb \\
 s.t. \quad Ax = b & s.t. \quad wA \leq c \\
 & x \geq 0
 \end{array}$$

After introducing artificial variables in (P) , we obtain the auxiliary problem

$$\begin{array}{ll}
 \min g = 1x_a \\
 s.t. \quad Ax + x_a = b \quad (b \geq 0) \\
 x, x_a \geq 0
 \end{array}$$

If a feasible solution \bar{w} to (D) is known, to maintain complementary slackness ($x(\bar{w}A - c) = 0$), let $x_j = 0$ (when $\bar{w}a_j \neq c_j$), so that the auxiliary problem is

$$\begin{array}{ll}
 (P') \min g = 1x_a \\
 s.t. \quad Ax + x_a = b \\
 x_j = 0, j \notin Q \\
 x_j \geq 0, j \in Q \\
 x_a \geq 0
 \end{array}$$

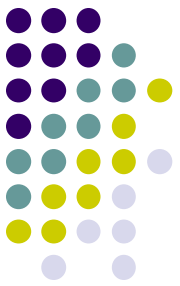
Wherein $Q = \{j \mid \bar{w}a_j = c_j, j = 1, \dots, n\}$. Problem (P') is called restrict problem corresponding to \bar{w} .

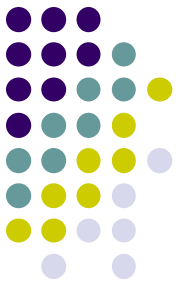
Solve (P') , obtain the optimal solution $\begin{pmatrix} x^* \\ x_a^* \end{pmatrix}$, it's a basic feasible solution to auxiliary problem of (P) .

If $x_a^* = 0$, x^* is the optimal solution to (P) (Because x^* and \bar{w} are feasible solutions to (P) and (D) , respectively. And complementary slackness is satisfied).

If $x_a^* \neq 0$, find another basic feasible solution \hat{w} to (D) to increase the objective function value of (D) .

Meanwhile, the optimal value corresponding to the limit problem of \hat{w} has relative reduction to that of \bar{w} (equivalent to eliminate artificial variables).





Consider the dual problem (D') to (P')

$$(D') \max v b$$

$$s.t. \quad v a_j \leq 0, j \in Q$$

$$v \leq 1$$

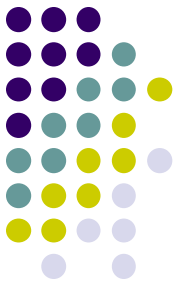
Let v^* be its optimal solution. If $v^* a_j \leq 0$ for any $j = 1, \dots, n$, v^* is still the optimal solution to the dual of auxiliary problem.

So $\begin{pmatrix} x^* \\ x_a^* \end{pmatrix}$ is the optimal solution to the auxiliary problem. Because of $x_a^* \neq 0$, the primal (P) has no optimal solution.

$$\text{Construct } \hat{w} = \bar{w} + \theta v^*, \text{ wherein } \theta = \min \left\{ -\frac{\bar{w} a_j - c_j}{v^* a_j} \mid v^* a_j > 0, j = 1, \dots, n \right\}.$$

We can prove that \hat{w} satisfies conditions. Meanwhile, $\begin{pmatrix} x^* \\ x_a^* \end{pmatrix}$ is a basic feasible solution to the

limit problem corresponding to \hat{w} , so take it as an initial solution to solve the limit problem corresponding to \hat{w} . Loop repeats.



1) Convert the programming problem into the form

$$\begin{aligned} \min \quad & cx \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

Find its dual initial solution w , which satisfies $wA \leq c$. (When $c \geq 0$, choose $w=0$ as the initial solution)

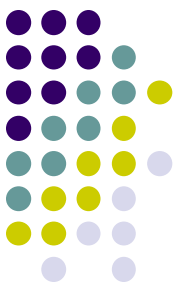
2) Let $Q = \{j \mid \bar{w}a_j = c_j, j = 1, \dots, n\}$. Solve the limit problem corresponding to w

$$\begin{aligned} \min \quad & g = 1x_a \\ \text{s.t.} \quad & Ax + x_a = b \\ & x_j = 0, j \notin Q \\ & x_j \geq 0, j \in Q \\ & x_a \geq 0 \end{aligned}$$

If its optimal value $g = 0$, terminates, the primal optimal solution is found; Otherwise, solve its dual and let the optimal solution be v .

3) If $vA \leq 0$, terminates, the primal has no optimal solution; Otherwise, let

$$\theta = \min \left\{ -\frac{\bar{w}a_j - c_j}{v^*a_j} \mid v^*a_j > 0, j = 1, \dots, n \right\}, \text{ and construct } w = w + \theta v^*, \text{ go to step 2.}$$



Dual initial solution

- Both dual simplex method and primal-dual simplex method need a dual feasible solution. Furthermore, dual simplex method requires a basic feasible solution.

For any LP problem, we can find a basis B by Gauss elimination method, and transform it into canonical form. If $\zeta_N \leq 0$, we have found a basic feasible solution $w = c_B B^{-1}$ to the dual. Otherwise, add a constraint $1x_N \leq M$, M is large enough constant. (So this constraint has no influence on the primal) It's called extended problem after adding constraint. (Provable: The primal has feasible problem \Leftrightarrow The extended problem has feasible solution, when M is large enough)

	z	x_B	x_N	x_{n+1}	RHS
z	1	0	ζ_N	0	z_0
x_B	0	I	\bar{N}	0	\bar{b}
x_{n+1}	0	0	1	1	M

Let $\zeta_k = \max \{\zeta_i\}$, pivot with pivot element $\bar{a}_{m+1,k}$ to get a new simplex tableau. Because of $\zeta \leq 0$, the new simplex tableau corresponds to a regular solution. Then we can solve the problem by dual simplex method or primal-dual simplex method.

There are two possible results:

- 1) The extended problem has no optimal solution \Rightarrow The primal has no optimal solution.
- 2) The extended problem has optimal solution $\bar{x} = \bar{x}_1 + \bar{x}_2 M$, the optimal value is $z_0 = z_1 + z_2 M$. Then the primal must have a feasible solution \hat{x} , and $z_1 + z_2 M \leq c\hat{x}$. Because M is large enough, there must be $z_2 \leq 0$. (Otherwise, $c\hat{x}$ can be arbitrary large.)
 - (i) $z_2 < 0$. When $M \rightarrow \infty, z_0 \rightarrow -\infty$. The problem is boundless.
 - (ii) $z_2 = 0$. $z_0 = z_1$ is the optimal value of the primal, $\bar{x} \geq 0$ is the optimal solution to the primal. If $\bar{x}_2 = 0$, \bar{x} is still a basic feasible solution. Otherwise, ($\bar{x}_2 \neq 0$), let $M_0 = \min \{M \mid \bar{x}_1 + \bar{x}_2 M \geq 0\}$, then $\bar{x}_0 = \bar{x}_1 + \bar{x}_2 M_0$ is also a basic feasible solution. And $\{x \mid x = \bar{x}_1 + \bar{x}_2 M, M \geq M_0\}$ represents a half-line face of the feasible region, every point in it is an optimal solution.